TESTS ON MEANS WITH ADDITIONAL INFORMATION

by

Morris L. Eaton[1] and Takeaki Kariya

Technical Report No. 243

February, 1975

University of Minnesota
Minneapolis, Minnesota

[1] This research was partially supported by a grant from the National Science Foundation - NSF-GP-34482.

Abstract

Let $X_1,\ldots,X_N$ be i.i.d. $N(\mu, \Sigma)$ with $X_i : p \times 1$, and partition $\mu = \binom{\mu_1}{\mu_2}$, $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$, with $\mu_i : p_i \times 1$, $\Sigma_{ij} : p_i \times p_j$, $i,j = 1,2$ where $p_1 + p_2 = p$. Suppose $W_1,\ldots,W_M$ are also i.i.d. $N(\mu_2, \Sigma_{22})$, $W_j : p_2 \times 1$. For the problem of testing $H_0 : \mu_2 = 0$ versus $H_1 : \mu_2 \neq 0$, the likelihood ratio test is derived and is shown to be uniformly most powerful invariant. For the problem of testing $\tilde{H}_0 : \mu = 0$ versus $\tilde{H}_1 : \mu \neq 0$, the likelihood ratio test is calculated and its distribution is given. It is also shown that, for $\tilde{H}_0$, a locally most powerful invariant test does not exist.

AMS subject classifications: Primary 62H15, secondary 62F05.

Key words and phrases: Tests on means, multivariate normal distribution, additional information, missing data, multivariate analysis of variance.

§ 1. Introduction and summary

Suppose $X_1, \ldots, X_N$ is a random sample from a p-dimensional normal distribution, $N(\mu, \Sigma)$ where μ and Σ are unspecified. Partition μ and Σ as

$$(1.1) \qquad \begin{cases} \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \\ \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \end{cases}$$

with $\mu_i : p_i \times 1$, $\Sigma_{ij} : p_i \times p_j$, $i,j = 1,2$ and $p_1 + p_2 = p$. Also, suppose that $W_1, \ldots, W_M$ is a random sample with $W_j : p_2 \times 1$ being distributed as $N(\mu_2, \Sigma_{22})$. The purpose of this paper is to investigate the effect of the extra data $(W_1, \ldots, W_M)$ on problems concerned with tests about the mean vector μ.

In Section 2, we consider the problem of testing $H_0 : \mu_2 = 0$ versus $H_1 : \mu_2 \neq 0$. For this problem, it is shown that the likelihood ratio test is uniformly most powerful within the class of invariant tests (UMPI). However, the standard arguments which establish that a test is UMPI cannot be employed for the current problem since an analytically tractable maximal invariant in the sample space is not available. The argument used in Section 2 relies on a representation theorem due to Wijsman (1967).

The problem of testing $H_0 : \mu \neq 0$ versus $H_1 : \mu \neq 0$ is considered in Section 3. For this problem we derive the likelihood ratio test and its distribution. Further, it is shown that a locally most powerful invariant test does not exist. Again, Wijsman's (1967) representation theorem is used.

We close the paper with a discussion of the results of this paper as compared with those in Eaton and Kariya (1974). Other problems involving partial extra information are also discussed.

§ 2. Testing $\mu_2 = 0$.

To treat the problem of testing $H_0 : \mu_2 = 0$ versus $H_1 : \mu_2 \neq 0$, we first reduce the data given in Section 1 by sufficiency. Let

$$S = \sum_{i=1}^{N} (X_i - \bar{X})(X_i - \bar{X})', \quad Z = \sqrt{M}\,\bar{W}, \quad Y = \sqrt{N}\,\bar{X},$$

$$V = \sum_{i=1}^{M} (W_i - \bar{W})(W_i - \bar{W})', \quad \nu = \sqrt{N}\,\mu$$

and $k = \sqrt{M}/\sqrt{N}$. Clearly (Y,Z,S,V) is a sufficient statistic and Y,Z,S,V are all independent. Further

$$(2.1) \qquad \begin{cases} Y \sim N(\nu, \Sigma) \\ Z \sim N(k\nu_2, \Sigma_{22}) \\ S \sim W(\Sigma, p, n) \\ V \sim W(\Sigma_{22}, p_2, m) \end{cases}$$

where $n = N-1$, $m = M-1$, $\nu = \binom{\nu_1}{\nu_2}$ is partitioned as μ, "$\sim$" means "is distributed", and $W(\Sigma, p, n)$ denotes the $p \times p$ Wishart distribution with n degrees of freedom and expectation $n\Sigma$.

The calculation of maximum likelihood estimators under H_0 and H_1 is somewhat simplified by writing (2.1) in the following conditional form:

$$
(2.2)\quad \begin{cases}
Y_1|Y_2 & \sim N(\nu_1 + B(Y_2 - \nu_2), \Sigma_{11.2}) \\
Y_2 & \sim N(\nu_2, \Sigma_{22}) \\
Z & \sim N(k\nu_2, \Sigma_{22}) \\
S_{11.2} & \sim W(\Sigma_{11.2}, p_1, n-p_2) \\
S_{12}|S_{22} & \sim N(BS_{22}, \Sigma_{11.2} \otimes S_{22}) \\
S_{22} & \sim W(\Sigma_n, p_2, n) \\
V & \sim W(\Sigma_{22}, p_2, m)
\end{cases}
$$

where $\Sigma_{11.2} \equiv \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$, $S_{11.2} = S_{11} - S_{12}S_{22}^{-1}S_{21}$, $B = \Sigma_{12}\Sigma_{22}^{-1}$, $\otimes$ denotes the Kronecker product, and $S_{12}|S_{22}$ denotes the conditional distribution of S_{12} given S_{22} (see Eaton (1972)).

Under H_0, it is not hard to show that the maximum likelihood estimates of the parameters are given by

$$
(2.3)\quad \begin{cases}
(N+M)\hat{\Sigma}_{22} = S_{22} + V + Y_2Y_2' + ZZ' \\
\hat{B} = S_{21}\, S_{22}^{-1} \\
N\hat{\Sigma}_{11.2} = S_{11.2} \\
\hat{\nu}_1 = Y_1 - \hat{B}\, Y_2 \, .
\end{cases}
$$

Under H_1, the maximum likelihood estimates of the parameters are given by

$$
(2.4)\quad \begin{cases}
\hat{\hat{\nu}}_2 = (Y_2 + kZ)/(1+k^2) \\
\hat{\hat{B}} = S_{12}S_{22}^{-1} \\
(N+M)\hat{\hat{\Sigma}}_{22} = S_{22} + V + (Y_2 - \hat{\hat{\nu}}_2)(Y_2 - \hat{\hat{\nu}}_2)' + (Z - k\hat{\hat{\nu}}_2)(Z - k\hat{\hat{\nu}}_2)' \\
N\hat{\hat{\Sigma}}_{11.2} = S_{11.2} \\
\hat{\hat{\nu}}_1 = Y_1 - \hat{\hat{B}}(Y_2 - \hat{\hat{\nu}}_2) \, .
\end{cases}
$$

Substituting these quantities into the likelihood function of the observations in (2.2), the likelihood ratio statistic is given by

$$(2.5) \quad \lambda = |\hat{\Sigma}_{22}|^{-\frac{N+M}{2}} / |\hat{\hat{\Sigma}}_{22}|^{-\frac{N+M}{2}} = \frac{|S_{22} + V + (Y_2 - \hat{\hat{\nu}}_2)(Y_2 - \hat{\hat{\nu}}_2)' + (Z - k\hat{\hat{\nu}}_2)(Z - k\hat{\hat{\nu}}_2)|^{\frac{N+M}{2}}}{|S_{22} + V + Y_2Y_2' + ZZ'|^{\frac{N+M}{2}}}$$

so

$$(2.6) \quad \lambda^{\frac{2}{N+M}} = \frac{|S_{22} + V + U(I-Q)U'|}{|S_{22} + V + UU'|}$$

where $U = (Y_2, Z) : p_2 \times 2$ and

$$Q : 2 \times 2 = \frac{1}{1+k^2} \begin{pmatrix} 1 & k \\ k & k^2 \end{pmatrix}$$

is an orthogonal projection of rank 1. Thus $I-Q : 2 \times 2$ is also an orthogonal projection of rank 1. Since $Q = aa'$ where

$$a = \frac{1}{\sqrt{1+k^2}} \begin{pmatrix} 1 \\ k \end{pmatrix} : 2 \times 1, \quad (I-Q) = bb' \quad \text{where} \quad b = \frac{1}{\sqrt{1+k^2}} \begin{pmatrix} -k \\ 1 \end{pmatrix}. \quad \text{Thus,}$$

$$(2.7) \quad \lambda^{\frac{2}{N+M}} = \frac{|S_{22} + V + Ubb'U'|}{|S_{22} + V + Uaa'U + Ubb'U'|}$$

$$= \frac{1}{1 + (Ua)'(S_{22} + V + Ubb'U')^{-1}(Ua)},$$

so the likelihood ratio test is equivalent to the test which rejects for large values of the statistic

(2.8) $F_0 = (Ua)'(S_{22} + V + Ubb'U')^{-1}(Ua)$.

Since $a'b = 0$, Ua and Ub are independent, $Ub \sim N(0, \Sigma_{22})$ so $S_{22} + V + Ubb'U \sim W(\Sigma_{22}, p_2, n+m+1)$ and $Ua \sim N(\sqrt{1+k^2}\,\nu_2, \Sigma_{22})$. From these facts it follows immediately that

$$\left(\frac{n+m+1-p_2}{p_2}\right) F_0$$

has a non-central F distribution - say $\mathfrak{F}_{p_2, n+m+1-p_2}(\tau)$ - with p_2 and $n+m+1-p_2$ degrees of freedom and non-centrality parameter $\tau = (1+k^2)\nu_2' \Sigma_{22}^{-1} \nu_2$. Thus, under H_0, standard tables of the F-distribution can be used to calculate percentage points of the distribution of F_0. This completes our discussion of the likelihood ratio test for H_0 versus H_1.

We now turn to the question of optimal properties of the test just derived. For the current discussion, the data in the form (2.1) will be analyzed. Let $A = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}$ be a $p \times p$ matrix with $A_{ii}: p_i \times p_i$ for $i = 1,2$ and $\det(A_{ii}) \neq 0$. For $b: p_1 \times 1$, the testing problem $H_0: \mu_2 = 0$ versus $H_1: \mu_2 \neq 0$, is invariant under the transformations on the sample space given by

$$(2.9) \qquad \begin{cases} Y \to AY + \binom{b}{0} \\ Z \to A_{22}Z \\ S \to ASA' \\ V \to A_{22}VA_{22}' \end{cases}$$

Formally, the testing problem is invariant under the group $G = \{g \mid g = (A,b)\}$ with A and b described above. Composition in the group G is defined by

$$(2.10)\qquad (A,b)\circ(\tilde{A},\tilde{b}) = (A\tilde{A},\ A_{11}\tilde{b} + b)\ .$$

The group action on the parameter space is

$$(2.11)\qquad \begin{aligned} &\mu \to A\mu + \binom{b}{0} \\ &\Sigma \to A\,\Sigma\,A'\ . \end{aligned}$$

The reader is referred to Lehmann (1959) or Eaton (1972) for a more complete discussion of decision problems invariant under groups of transformation.

Proposition 2.12: $\delta = \mu_2' \Sigma_{22}^{-1} \mu_2$ is a maximal invariant parameter under the group action given by (2.11).

Proof: The proof is routine and is omitted.

Now, let $\mathcal{D}_\alpha$ be the class of level α test functions which are invariant under the group G. The remainder of this section is devoted to showing that the likelihood ratio test derived earlier is a uniformly most powerful test in $\mathcal{D}_\alpha$. Ordinarily, the proof of such a result would proceed along the following lines:

(i) First, compute a nice function of the data, say $T = T(Y,Z,S,V)$, which is maximal invariant under the group action of (2.9).

(ii) Next, derive the distribution of T, which will depend only on the parameter δ, say P_δ^T.

(iii) On the basis of the analytical form of $R_\delta(t) \equiv \frac{dP_\delta^T}{dP_0^T}(t)$, argue that a UMP test of $H_0^*: \delta = 0$ versus $H_1^*: \delta > 0$ exists and then show that this UMP test is equivalent to the likelihood ratio test.

However, for the problem at hand, an analytically tractable maximal invariant seems difficult to find so that we have not been able to carry out (i) of the above proceedure. An alternative to the above proceedure is to use a result due to Wijsman (1967) to calculate the form of R_δ directly without going through steps (i) and (ii). Now, we proceed with the details.

Since the testing problem is invariant under G, to analyze the power functions of tests $\phi \in \mathcal{D}_\alpha$, we can take $\Sigma = I$ without loss of generality. Now, write

$$(2.13)\quad \begin{cases} S \sim uu', & u \sim N(0, I_p \otimes I_p) \\ V \sim vv', & v \sim N(0, I_{p_2} \otimes I_m) \end{cases}$$

where u is $n \times p$ and v is $m \times p_2$. Then we have the data

$$(2.14)\quad \begin{cases} Y \sim N(\nu, I_p) \\ Z \sim N(k\nu_2, I_{p_2}) \\ u \sim N(0, I_p \otimes I_n) \\ v \sim N(0, I_{p_2} \otimes I_m) \end{cases}$$

and the action of G is

$$(2.15)\quad \begin{cases} Y \to AY + b \\ Z \to A_{22} Z \\ u \to Au \\ v \to A_{22} v . \end{cases}$$

The joint density of (y,z,u,v) with respect to Lebesgue measure is

$$(2.16)\qquad p_\nu(y,z,u,v) = c\,\exp[-\tfrac{1}{2}(y-\nu)'(y-\nu) -\tfrac{1}{2}(z-k\nu_2)'(z-k\nu_2) - \tfrac{1}{2}\mathrm{tr}\,uu' - \tfrac{1}{2}\mathrm{tr}\,vv']$$

where c is a constant.

<u>Proposition 2.17:</u> A left invariant measure on G is given by

$$(2.18)\qquad \gamma(dA,db) = |A_{11}A_{11}'|^{-\frac{p_1+p_2+1}{2}}\,|A_{22}A_{22}'|^{-p_2/2}\,dA_{11}dA_{12}dA_{22}db\ .$$

<u>Proof:</u> The proof is routine and is omitted.

The Jacobian of the transformation (2.15) is, for $g = (A,b)$,

$$(2.19)\qquad J(g) = |AA'|^{-\frac{n+1}{2}}\,|A_{22}A_{22}'|^{-\frac{m+1}{2}}\ .$$

Let T be any maximal invariant function of the data and let P_δ^T denote its distribution. According to Wijsman's (1967) result, the Radon-Nikodym derivative $R_\delta \equiv dP_\delta^T/dP_0^T$ is given by

$$(2.20)\qquad R_\delta = \frac{\int_G p_\nu(g(y,z,u,v))J(g)^{-1}\gamma(dg)}{\int_G p_0(g(y,z,u,v))J(g)^{-1}\gamma(dg)}\ .$$

<u>Theorem 2.21:</u> R_δ is given by

$$(2.22)\qquad R_\delta = H(\delta)\sum_{j=0}^{\infty} c_j\delta^j[(kz+y_2)'(S_{22}+V+UU')^{-1}(kz+y_2)]^j$$

where

$$c_j = \frac{2^j}{j!}\,\frac{\Gamma(\frac{p_2}{2})\Gamma(j+\frac{1}{2})\Gamma(\frac{m+n+2}{2}+j)}{\Gamma(\frac{1}{2})\Gamma(j+p_2/2)\Gamma(\frac{m+n+2}{2})}\ ,$$

U is defined in (2.6), and $H(\delta) = \exp[-\frac{1}{2}(1+k^2)\delta]$.

Proof: First note that

$$(2.23)\qquad \mathrm{tr}Auu'A' = \mathrm{tr}ASA' = \mathrm{tr}A_{11}S_{11.2}A_{11}' + \mathrm{tr}A_{22}S_{22}A_{22}' + \mathrm{tr}(A_{12}S_{22}^{\frac{1}{2}} + A_{11}S_{12}S_{22}^{-\frac{1}{2}})(A_{12}S_{22}^{\frac{1}{2}} + A_{11}S_{12}S_{22}^{-\frac{1}{2}})' ,$$

where $S_{22}^{\frac{1}{2}}$ is a square root of S_{22}. Thus, since $\delta = \nu_2'\nu_2$ as $\Sigma_{22} = I$,

$$(2.24)\qquad p_\nu(g(y,z,u,v)) = c\exp[-\tfrac{1}{2}(Ay + \tbinom{b}{0} - \nu)'(Ay + \tbinom{b}{0} - \nu)$$
$$-\tfrac{1}{2}(A_{22}z-k\nu_2)'(A_{22}z-k\nu_2) - \tfrac{1}{2}\mathrm{tr}Auu'A' - \tfrac{1}{2}\mathrm{tr}A_{22}vv'A_{22}']$$
$$= c\,H(\delta)\exp[-\tfrac{1}{2}(A_{11}y_1 + A_{12}y_2 + b-\nu_1)'(A_{11}y_1 + A_{12}y_2 + b-\nu_1)$$
$$+ \nu_2'A_{22}(kz+y) - \tfrac{1}{2}\mathrm{tr}A_{22}(V + y_2y_2' + zz')A_{22}'$$
$$-\tfrac{1}{2}\mathrm{tr}\,A_{11}S_{11.2}A_{11}' - \tfrac{1}{2}\mathrm{tr}\,A_{22}S_{22}A_{22}'$$
$$-\tfrac{1}{2}\mathrm{tr}(A_{12}S_{22}^{\frac{1}{2}} + A_{11}S_{12}S_{22}^{-\frac{1}{2}})(A_{12}S_{22}^{\frac{1}{2}} + A_{11}S_{12}S_{22}^{-\frac{1}{2}})'] .$$

Integrating first over b, then over A_{12} and next over A_{11}, we see that all these factors cancel in the expression for R_δ. Thus, setting $W = S_{22} + V + y_2y_2' + zz'$,

$$(2.25)\qquad R_\delta = \frac{H(\delta)\int \exp[\nu_2'A_{22}(kz+y_2) - \frac{1}{2}\mathrm{tr}\,A_{22}WA_{22}']\,|A_{22}A_{22}'|^{\frac{m+n+2-p_2}{2}}\,dA_{22}}{\int \exp[-\frac{1}{2}\mathrm{tr}\,A_{22}WA_{22}']\,|A_{22}A_{22}'|^{\frac{m+n+2-p_2}{2}}\,dA_{22}}$$
$$= H(\delta)\,\frac{\int \exp[\nu_2'A_{22}W^{-\frac{1}{2}}(kz+y_2) - \frac{1}{2}\mathrm{tr}\,A_{22}A_{22}']\,|A_{22}A_{22}'|^{\frac{m+n+2-p_2}{2}}\,dA_{22}}{\int \exp[-\frac{1}{2}\mathrm{tr}\,A_{22}A_{22}']\,|A_{22}A_{22}'|^{\frac{m+n+2-p_2}{2}}\,dA_{22}} .$$

The expression (2.22) for R_δ now follows immediately from Lemma 1 in the appendix, since $W = S_{22} + V + UU'$. This completes the proof.

Theorem 2.26: A uniformly most powerful test in $\mathcal{D}_\alpha$ of $H_0^*: \delta = 0$ versus $H_1^*: \delta > 0$ is given by

$$(2.27)\quad \phi^* = \begin{cases} 1 & \text{if } (kz+y_2)'(S_{22} + V + UU')^{-1}(kz+y_2) \geq k_0 \\ 0 & \text{otherwise} \end{cases}$$

where k_0 is chosen so that $\phi^* \in \mathcal{D}_\alpha$.

Proof: Since R_δ is an increasing function of $(kz+y_2)'(S_{22}+V+UU')^{-1}(kz+y_2)$, the result follows immediately from the Neyman-Pearson Lemma.

Theorem 2.28: The test ϕ^* is equivalent to the likelihood ratio test.

Proof: As in equation (2.7), write $UU' = Uaa'U' + Ubb'U'$ and note that $Ua = \frac{1}{\sqrt{1+k^2}}(kz+y_2)$. Thus

$$(2.29)\quad (kz+y_2)'(S_{22} + V + UU')^{-1}(kz+y_2) =$$

$$(1+k^2)(Ua)'(S_{22} + V + Ubb'U' + Uaa'U')^{-1}(Ua)$$

$$= (1+k^2)\,\frac{(Ua)'(S_{22} + V + Ubb'U)^{-1}(Ua)}{1 + (Ua)'(S_{22} + V + Ubb'U')^{-1}(Ua)}\,.$$

The equivalence of ϕ^* and the likelihood ratio test is now clear.

§ 3. Testing that $\nu_1 = 0$ and $\nu_2 = 0$.

In this section we derive the likelihood ratio test for testing $H_0: \nu = 0$ versus $H_1: \nu \neq 0$ $(\nu = \binom{\nu_1}{\nu_2})$ and find the distribution of the likelihood ratio test under H_0. In addition, we outline an argument which shows that a locally most powerful invariant test of H_0 versus H_1 does not exist. Again, Wijsman's (1967) theorem is used.

To derive the likelihood ratio test, the data is again considered in the conditional form:

$$(3.1)\quad \left\{\begin{array}{rcl} Y_1|Y_2 & \sim & N(\nu_1+B(Y_2-\nu_2),\Sigma_{11.2}) \\ Y_2 & \sim & N(\nu_2,\Sigma_{22}) \\ Z & \sim & N(k\nu_2,\Sigma_{22}) \\ S_{11.2} & \sim & W(\Sigma_{11.2},p_1,n-p_2) \\ S_{12}|S_{22} & \sim & N(BS_{22},\Sigma_{11.2}\otimes S_{22}) \\ S_{22} & \sim & W(\Sigma_{22},p_2,n) \\ V & \sim & W(\Sigma_{22},p_2,m) \end{array}\right.$$

Under $H_0: \nu_1 = 0, \quad \nu_2 = 0,$ the maximum likelihood estimates of the remaining parameters are given by

$$(3.2)\quad \left\{\begin{array}{rl} (N+M)\hat{\Sigma}_{22} & = S_{22}+V+Y_2Y_2'+ZZ' \\ \hat{B} & = (S_{12}+Y_1Y_2')(S_{22}+Y_2Y_2')^{-1} \\ N\hat{\Sigma}_{11.2} & = (S_{11}+Y_1Y_1')-(S_{12}+Y_1Y_2')(S_{22}+Y_2Y_2')^{-1}(S_{12}+Y_1Y_2'). \end{array}\right.$$

Under $H_1: \nu \neq 0,$ the maximum likelihood estimates of the parameters are

$$(3.3)\quad \left\{\begin{array}{rl} \hat{\hat{\nu}}_2 & = (Y_2+kZ)/(1+k^2) \\ \hat{\hat{B}} & = S_{12}S_{22}^{-1} \\ (N+M)\hat{\hat{\Sigma}}_{22} & = S_{22}+V+(Y_2-\hat{\hat{\nu}}_2)(Y_2-\hat{\hat{\nu}}_2)'+(Z-k\hat{\hat{\nu}}_2)(Z-k\hat{\hat{\nu}}_2)' \\ N\hat{\hat{\Sigma}}_{11.2} & = S_{11.2} \\ \hat{\hat{\nu}}_1 & = Y_1-\hat{\hat{B}}(Y_2-\hat{\hat{\nu}}_2)\ . \end{array}\right.$$

Substituting these values into the likelihood function of the data in (3.1) under H_0 and H_1 and forming the ratio, the likelihood ratio statistic is

$$(3.4)\qquad \lambda = \frac{|\hat{\Sigma}_{11.2}|^{-\frac{N}{2}}|\hat{\Sigma}_{22}|^{-\frac{M+N}{2}}}{|\hat{\hat{\Sigma}}_{11.2}|^{-\frac{N}{2}}|\hat{\hat{\Sigma}}_{22}|^{-\frac{M+N}{2}}} \equiv \lambda_1\lambda_2$$

where

$$(3.5)\qquad \begin{cases} \lambda_1^{\frac{2}{N}} = \dfrac{|\hat{\hat{\Sigma}}_{11.2}|}{|\hat{\Sigma}_{11.2}|} = \dfrac{|S_{11.2}|}{|S_{11}+ Y_1Y_1' - (S_{12}+Y_1Y_2')(S_{22}+Y_2Y_2')^{-1}(S_{12}Y_1Y_2')|} \\[2ex] \lambda_2^{\frac{2}{N+M}} = \dfrac{|\hat{\hat{\Sigma}}_{22}|}{|\hat{\Sigma}_{22}|} \,. \end{cases}$$

<u>Theorem 3.6:</u> Under H_0, λ_1 and λ_2 are independent. Further, under H_0,

$$(3.7)\qquad \lambda_1^{\frac{2}{N}} \sim \frac{1}{1 + F_1}$$

and

$$(3.8)\qquad \lambda_2^{\frac{2}{N+M}} \sim \frac{1}{1 + F_2}$$

where $(\frac{n+m+1-p_2}{p_2})F_2$ has an $\mathfrak{F}_{p_2,n+m+1-p_2}$ distribution and $(\frac{n+1-p}{p_1})F_1$ has an $\mathfrak{F}_{p_1,n+1-p}$ distribution.

<u>Proof:</u> First, write

$$(3.9)\qquad \lambda_1^{\frac{2}{N}} = \frac{|S|}{|S + YY'|}\,\frac{|S_{22}+ Y_2Y_2'|}{|S_{22}|}$$

$$= \frac{1 + Y_2'S_{22}^{-1}Y_2}{1 + YS^{-1}Y}$$

$$= \frac{1 + Y_2' S_{22}^{-1} Y_2}{1 + (Y_1 - \hat{B}Y_2)' S_{11.2}^{-1} (Y_1 - \hat{B}Y_2) + Y_2' S_{22}^{-1} Y_2}$$

$$= \frac{1}{1 + \dfrac{(Y_1 - \hat{B}Y_2)' S_{11.2}^{-1} (Y_1 - \hat{B}Y_2)}{1 + Y_2' S_{22}^{-1} Y_2}} .$$

Conditional on (S_{22}, Y_2), $Y_1 \sim N(BY_2, \Sigma_{11.2})$ and $\hat{B}Y_2 = S_{12} S_{22}^{-1} Y_2 \sim N(BY_2, Y_2' S_{22}^{-1} Y_2 \Sigma_{11.2})$. Thus under H_0 and conditional on (S_{22}, Y_2),

$$(3.10) \qquad \frac{(Y_1 - \hat{B}Y_2)}{\sqrt{1 + Y_2' S_{22}^{-1} Y_2}} \sim N(0, \Sigma_{11.2}) .$$

Since $S_{11.2} \sim W(\Sigma_{11.2}, p_1, n-p_2)$, the conditional distribution of $F_1 = (Y_1 - \hat{B}Y_2)' S_{11.2}^{-1} (Y_1 - \hat{B}Y_2)/(1 + Y_2' S_{22}^{-1} Y_2)$ is $p_1/(n+1-p)\ \mathfrak{F}_{p_1, n+1-p_1}$. Hence F_1 has the same distribution unconditionally. Since F_1 and $\lambda_2^{\frac{2}{N+M}}$ are clearly conditionally independent given (S_{22}, Y_2), they are, under H_0, unconditionally independent since the distribution of F_1 does not depend on the conditioned variables. The distribution of $\lambda_2^{\frac{2}{N+M}}$ follows as in (2.7) and the discussion following (2.7). This completes the proof.

From the above result, it follows that under H_0, λ is distributed as the product of two independent random variables which are each powers of Beta random variables with the powers being different. Unfortunately, existing tables do not allow the exact calculation of the percentage points of λ under H_0.

The following interpretation of λ is interesting. λ_2 is the likelihood ratio statistic for testing $H_0: \nu_2 = 0$ versus $H_1: \nu_2 \neq 0$ where λ_1 is the likelihood ratio statistics for testing $H_0^*: \nu_1 = 0$, $\nu_2 = 0$ versus $H_1^*: \nu_1 \neq 0$, $\nu_2 = 0$. Thus λ can be interpreted as first testing H_0 and after accepting H_0, testing H_0^* under the assumption that H_0 is true. This type of interpretation and decomposition of likelihood ratio statistics for testing normal means occurs in other contexts. For example, see Hogg(1961) for a univariate normal example, Eaton (1972) for the multivariate analysis of variance case, and Kariya (1974) for the application of these ideas to the multivariate linear growth curve model.

In contrast to the situation in Eaton and Karyia (1974), we now indicate that there does not exist a locally most powerful invariant test for testing $H_0: \nu = 0$ versus $H_1: \nu \neq 0$. Consider the data in the form (2.1):

$$(3.11) \qquad \begin{cases} Y \sim N(\nu, \Sigma) \\ Z \sim N(k\nu_2, \Sigma_{22}) \\ S \sim W(\Sigma, p, n) \\ V \sim W(\Sigma_{22}, p_2, m) \;. \end{cases}$$

The testing problem is invariant under the group $G_0 = \{A \mid A = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}$, $\det A_{ii} \neq 0$, $i = 1,2\}$ as described in Section 2. The action of G_0 on the sample space is given by (2.9) (with $b = 0$) and the action on the parameter space is given by (2.11) (with $b = 0$). The following is clear.

Proposition 3.12: A maximal invariant in the parameter space is

$$(3.13) \qquad \delta = \binom{\delta_1}{\delta_2} = \begin{bmatrix} (\nu_1 - \Sigma_{12}\Sigma_{22}^{-1}\nu_2)'\Sigma_{11.2}^{-1}(\nu_1 - \Sigma_{12}\Sigma_{22}^{-1}\nu_2) \\ \nu_2'\,\Sigma_{22}^{-1}\,\nu_2 \end{bmatrix} .$$

As with the testing problem in Section 2, an analytically tractable maximal invariant (under G_0) seems difficult to find. Let P_δ^T denote the probability distribution of any maximal invariant T at the parameter point $\delta = \binom{\delta_1}{\delta_2}$. Further, set $W = S + YY'$ and $X = V + ZZ'$.

Theorem 3.14: The Radon-Nikodym derivative

$$(3.15) \qquad R_\delta = dP_\delta^T / dP_0^T = H(\delta)F_1F_2F_3$$

where

$$(3.16) \qquad \begin{cases} H(\delta) = \exp[-\tfrac{1}{2}\delta_1 - \tfrac{1}{2}(1+k^2)\delta_2] \\ F_1 \;\; = \sum_{j=0}^{\infty} \delta_1^j (Y_1 - W_{12}W_{22}^{-1}Y_2)'W_{11.2}^{-1}(Y_1 - W_{12}W_{22}^{-1}Y_2)c_j^0 \\ F_2 \;\; = \sum_{j=0}^{\infty} \delta_2^j (Y_2 + kZ)'(W_{22} + X)^{-1}(Y_2 + kZ)d_j^0 \\ F_3 \;\; = \exp[\tfrac{1}{2}\,\delta_1 Y_2'W_{22}^{-1}Y_2] \end{cases}$$

and

$$(3.17) \qquad \begin{cases} c_j^0 = \dfrac{1}{j!}\,\dfrac{\Gamma(\frac{p_1}{2})\Gamma(j+\frac{1}{2})}{\Gamma(\frac{1}{2})\Gamma(j+\frac{p_1}{2})}\,\dfrac{2^j\,\Gamma(\frac{n+1-p_2}{2}+j)}{\Gamma(\frac{n+1-p_2}{2})} \\[2ex] d_j^0 = \dfrac{1}{j!}\,\dfrac{\Gamma(\frac{p_2}{2})\Gamma(j+\frac{1}{2})}{\Gamma(\frac{1}{2})\Gamma(j+\frac{p_2}{2})}\,\dfrac{2^j\,\Gamma(\frac{m+n+2}{2}+j)}{\Gamma(\frac{m+n+2}{2})} \end{cases} .$$

Proof: The proof of this theorem is similar to the proof of Theorem 2.21. The details are omitted.

Corollary 3.18: For δ_1 and δ_2 small,

$$(3.19)\qquad R_\delta = 1 + \delta_1(\xi_1 + \xi_3 - \tfrac{1}{2}) + \delta_2(\xi_2 - \tfrac{1}{2}(1+k^2)) + o(\delta_1 + \delta_2)$$

where

$$(3.20)\qquad \begin{cases} \xi_1 = c_1^0(Y_1 - W_{12}W_{22}^{-1}Y_2)'W_{11.2}^{-1}(Y_1 - W_{12}W_{22}^{-1}Y_2) \\ \xi_2 = d_1^0(Y_2 + kZ)(W_{22} + X)^{-1}(Y_2 + kZ) \\ \xi_3 = \tfrac{1}{2}\, Y_2'W_{22}^{-1}Y_2 \end{cases}$$

and the remainder term $o(\delta_1 + \delta_2)$ is uniform in (Y,S,V,Z).

Proof: This follows immediately from (3.16) by simply multiplying out the constant and linear terms in $H(\delta)$ and F_i, $i = 1,2,3$ and noting that the error term is uniform in (Y,S,V,Z).

Now, let φ be a level α test function of $H_0: \nu = 0$ versus $H_1: \nu \neq 0$ which is invariant under the group G_0. From (3.19), the power function of φ can be expressed by

$$(3.21)\qquad \mathcal{E}_{\delta\varphi} = \alpha + \mathcal{E}_\delta[\varphi\{\delta_1(\xi_1 + \xi_3 - \tfrac{1}{2}) + \delta_2(\xi_2 - \tfrac{1}{2}(1+b^2))\}] + o(\delta_1 + \delta_2)$$

where the remainder term $o(\delta_1 + \delta_2)$ is uniform in φ. Using the Gerneralized Neyman-Pearson Lemma (Lehmann (1959)), it follows from (3.21) that a locally most powerful test for alternatives of the form $\delta_2 = \gamma\delta_1 > 0$ (with γ a known positive constant), exists and it clearly depends on γ. Hence a locally most powerful invariant test does not exist for $H_0: \delta = 0$ versus $H_1: \delta_1 > 0$, $\delta_2 > 0$. This completes our discussion of locally most powerful test of H_0.

§ 4. Discussion.

In an earlier paper, Eaton and Kariya (1974) investigated the effect of having additional observations on certain coordinates when testing for independence in multivariate normal populations. In this situation, the likelihood ratio test completely ignores the extra observations. However, a locally most powerful test does exist and involves the extra observations.

The purpose of the present paper was to investigate the effect of having extra observations on testing problems involving the mean for observations from the multivariate normal. As we have seen, the effect of the extra observations depends very much on the particular hypothesis involving the mean vector.

The two problems above do have an interesting common feature. Namely, the calculation of a reasonable maximal invariant in the sample space does not seem possible. Thus, to investigate the power functions of invariant tests, Wijsman's (1967) Theorem was used.

It is clear that a variety of open questions remain as to the effect of extra observations on mean testing problems. One possible context in which to study such problems is the multivariate analysis of variance (MANOVA) context. One form for the MANOVA model is the following. Suppose $Y:N \times p \sim N(X_1B, I_N \times \Sigma)$ where Y is a matrix of observations, $X_1:N \times q$ is a known matrix of rank q, and $B:q \times p$ is a matrix of unknown parameters. Partition B as $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$ where $B_{i,j}$ is $q_i \times p_j$, $i,j = 1,2$ and $p_1 + p_2 = p$, $q_1 + q_2 = q$. Consider "extra observations" $Z:M \times p_2 \sim N(X_2B_{22}, I_M \times \Sigma_{22})$ where $X_2:M \times q_2$ is a known matrix of rank q_2. One of the standard problems in MANOVA is to test $H_0:CB = 0$ versus $H_1:CB \neq 0$, where C is $r \times q$ of rank r. A central problem is to discover how the presence of Z effects the testing problem. For example, under what conditions will a locally most powerful test exist? Or, under what conditions will locally minimax tests exist?

Appendix

Let $\mathcal{O}(r)$ denote the group of $r \times r$ orthogonal matrices, let $\mathcal{G}\ell(s)$ denote the group of $s \times s$ non-singular real matrices, and let $G_T^+(s)$ be the group of $s \times s$ lower triangular matrices with positive diagonal elements.

Lemma 1: If Γ is distributed uniformly on $\mathcal{O}(r)$, (i.e., Γ has invariant probability measure on $\mathcal{O}(r)$ as its distribution) then γ_{11}^2 has a beta distribution with parameters $\frac{1}{2}$ and $\frac{r-1}{2}$. Further,

$$\text{(A.1)} \qquad \mathcal{E}\, \gamma_{11}^{2k} = \frac{\Gamma(\frac{r}{2})\Gamma(\frac{2k+1}{2})}{\Gamma(\frac{1}{2})\Gamma(\frac{2k+r}{2})} .$$

Proof: Since Γ is uniform on $\mathcal{O}(r)$, the first row of Γ has the same distribution as $\frac{X}{\|X\|}$ where $X \sim N(0,I_r)$. Thus, $\gamma_{11}^2 \sim X_1^2/\sum_1^r X_i^2$. That $\gamma_{11}^2 \sim Be(\frac{1}{2}, \frac{r-1}{2})$ now follows from Wilks (1962) and (A.1) is an easy computation.

Lemma 2: Let $B \in \mathcal{G}\ell(s)$ have a density on $\mathcal{G}\ell(s)$ given by

$$\text{(A.2)} \quad p(B)\frac{dB}{|BB'|^{s/2}} = \frac{\exp[-\frac{1}{2}\text{tr } BB']|BB'|^{\frac{m-s}{2}} dB}{\int \exp[-\frac{1}{2}\text{tr } BB']|BB'|^{\frac{m-s}{2}} dB}$$

where m is an integer, $m \geq s$. Then

$$\text{(A.4)} \qquad \mathcal{E}\, e^{x'By} = \sum_{j=0}^{\infty} d_j (x'x)^j (y'y)^j$$

$$\text{where} \quad d_j = \frac{2^j}{j!} \frac{\Gamma(\frac{s}{2})\Gamma(j+\frac{1}{2})}{\Gamma(\frac{1}{2})\Gamma(j+\frac{s}{2})} \frac{\Gamma(\frac{m}{2}+j)}{\Gamma(\frac{m}{2})} .$$

Proof: First note that $dB/|BB'|^{\frac{s}{2}}$ is both right and left invariant measure on $G\ell(s)$. Each $B \in G\ell(s)$ can be written uniquely as $T\Gamma$ with $T \in G_T^+(p)$ and $\Gamma \in \mathcal{O}(s)$ (Eaton (1972)). Let $\nu_\ell(dT)$ be a left invariant measure on $G_T^+(s)$ and $\nu(d\Gamma)$ be invariant probability measure on $\mathcal{O}(s)$. It follows from the uniqueness of invariant measures that

(A.5) $$\int_{G\ell(s)} f(B) \frac{dB}{|BB'|^{s/2}} = k_1 \int_{G_T^+(s)} \int_{\mathcal{O}(s)} f(T\Gamma)\ \nu_\ell(dT)\nu(d\Gamma)$$

for all integrable f defined on $G\ell(s)$ where k_1 is a constant. Thus

(A.6) $$\int e^{x'By}\ p(B) \frac{dB}{|BB'|^{s/2}} = \int\int e^{x'T\Gamma y}\ p_1(T)\nu_\ell(dT)\nu(d\Gamma)$$

where

(A.7) $$p_1(T) = \frac{\exp[-\frac{1}{2}\mathrm{tr}\ TT']|TT'|^{\frac{m}{2}}}{\int \exp[-\frac{1}{2}\mathrm{tr}\ TT']|TT'|^{\frac{m}{2}}\ \nu_\ell(dT)} .$$

Now,

(A.8) $$\begin{aligned}\int e^{x'T\Gamma y}\ \nu(d\Gamma) &= \int \exp[x'TT'x)^{\frac{1}{2}}(y'y)^{\frac{1}{2}}\ \gamma_{11}]\nu(d\Gamma) \\ &= \int \sum_{j=0}^{\infty} \frac{1}{j!} (x'TT'x)^j(y'y)^j\ \gamma_{11}^{2j}\ \nu(d\Gamma) \\ &= \sum_{j=0}^{\infty} \frac{1}{j!} (x'TT'x)^j(y'y)^j \frac{\Gamma(\frac{s}{2})\Gamma(j+\frac{1}{2})}{\Gamma(\frac{1}{2})\Gamma(j+\frac{s}{2})} .\end{aligned}$$

The last step follows from (A.1).

Now, write $S = TT'$. Again, it follows from the uniqueness of invariant measures that

$$\text{(A.9)} \qquad \int_{S_p^+} g(S) \frac{dS}{|S|^{\frac{s+1}{2}}} = k_2 \int_{G_T^+(s)} g(TT')\nu_\ell(dT)$$

Where S_s^+ is the space of $s \times s$ positive definite symmetric matrices, g is an integrable function and k_2 is a constant. Hence

$$\text{(A.10)} \qquad \int (x'TT'x)^j \, p(T) \, \nu_\ell(dT) =$$

$$\frac{\int (x'Sx)^j \exp[-\tfrac{1}{2}\mathrm{tr}\, S]|S|^{\frac{m-p-1}{2}} dS}{\int \exp[-\tfrac{1}{2}\mathrm{tr}\, S]|S|^{\frac{m-p-1}{2}} dS} .$$

Thus $S \sim W(I,s,m)$ so $\frac{x'Sx}{x'x}$ has a χ_m^2-distribution. Therefore,

$$\text{(A.11)} \qquad \int (x'TT'x)^j \, p(T) \, \nu_\ell(T) = (x'x)^j \frac{2^j\Gamma(\frac{m}{2}+j)}{\Gamma(\frac{m}{2})} .$$

Combining (A.11) with (A.8) yields A.4. This completes the proof.

The above argument shows that when B has a density (A.2), then b_{11}^2 is distributed as the product of a $Be(\frac{1}{2}, \frac{s-1}{2})$ and an independent χ_m^2.

References

Eaton, M. L. (1972). Multivariate Statistical Analysis. Institute of Math. Statist. University of Copenhagen, Copenhagen, Denmark.

Eaton, M. L. and Kariya, T. (1974). "Testing for independence with additional information", University of Minnesota Technical Report No. 238.

Hogg, R. V. (1961). "On the resolution of statistical hypotheses". Jour. Amer. Statist. Ass'n., 56, pp. 978-89.

Kariya, T. (1974). "Hypothesis testing problems in the general MANOVA model, Ph.D. Thesis, University of Minnesota.

Lehmann, E. L. (1959). Testing Statistical Hypotheses. Wiley, New York.

Wijsman, R. A. (1967). "Cross-sections of orbits and their applications", Proc. Fifth Berk. Symp. Math. Statist. Prob., 1, 389-400, University of California.

Wilks, S. S. (1962). Mathematical Statistics, Wiley, New York.